

# Autoevaluación

1. **¿Cuál de las siguientes afirmaciones es correcta en relación a ETL e integración de datos?**

- a) Son equivalentes en todo caso.
- b) El proceso es el mismo pero en integración de datos el destino es un almacén de datos.
- c) Integración de datos es uno de los pasos que se realizan dentro del proceso de ETL.
- d) No son equivalentes.

2. La fase de extracción del proceso ETL consiste en extraer los metadatos incluidos dentro de los ficheros.

La fase de transformación es en la que se realiza todo el proceso de transformación de los datos para dejarlos en el formato finalmente deseado.

En la fase de carga de ETL es cuando se cargan datos desde disco a memoria. Falso

3. **¿En qué momento es mejor realizar la fase de extracción en un proceso de ETL?**

- a) Cualquier momento es bueno.
- b) Siempre por la noche.
- c) Si va interferir con el correcto funcionamiento de los sistemas que contienen las fuentes se realizará en el momento en el que tales sistemas o fuentes se encuentren menos cargados.
- d) Siempre por la mañana.

4. **¿Cuál de las siguientes no es un posible tipo de transformación a realizar con los datos en un proceso de ETL?**

- a) Realizar unificaciones de datos que vienen de distintas fuentes.
- b) Eliminar el contenido de una base de datos.
- c) Transponer filas en columnas o columnas en filas.
- d) Seleccionar sólo determinadas columnas o atributos de cada registro.

Las transformaciones en ETL no implican eliminar toda la base de datos.

5. **¿Qué les ocurre a los datos que ya hay en el almacenamiento destino de un ETL cuando llegan datos nuevos?**

- a) Se eliminan.
- b) Se actualizan.
- c) Depende el uso.
- d) En ocasiones se eliminan, en ocasiones no quedan afectados y en ocasiones son modificados.
- e) Depende el uso.
- f) En ocasiones se eliminan y en ocasiones son modificados.

6. Apache Sqoop es un software distribuido que permite obtener datos en streaming desde gran cantidad de fuentes no estructuradas o semiestructuradas. Falsa

Apache Sqoop es una herramienta de línea de comandos que nos permite obtener datos desde bases de datos relacionales.

Apache Sqoop permite enviar datos a HDFS.

Apache Flume permite enviar datos a HDFS.

7. ¿Cuáles de las siguientes son características propias de Apache Sqoop?

a) Se puede emplear mediante interfaz de línea de comandos.

b) Obtiene datos desde fuentes no estructuradas.

c) Tiene mecanismos para evitar sobrecargar las fuentes.

d) Permite acceso programático mediante JDBC.

8. ¿Cuáles de las siguientes son características propias de Apache Flume?

a) Permite conectar con cualquier fuente JDBC.

b) Alto ancho de banda y baja latencia.

c) Tolerancia a fallos.

d) Basado en eventos.

9. El proceso de Integración de Datos:

a) Consiste en una replicación modificada de los datos.

b) Busca entregar una visión unificada de los datos a aquellos usuarios o procesos que vayan a utilizarlos.

c) Es un paso previo al proceso de ETL.

d) Es una de las actividades englobadas dentro del proceso de ETL.

La integración de datos tiene como objetivo unificar y permitir acceso coherente a los datos.

10. ¿Cuál de los siguientes niveles de integración de datos es el que nos permite acceder a los datos mediante una visión unificada de los mismos?

a) Integración basada en aplicación.

b) Integración virtual.

c) Integración basada en *middleware*.

d) Integración manual.

11. El interesado es la persona física o jurídica, autoridad pública, servicio u otro organismo al que se comuniquen datos personales, se trate o no de un tercero.

El responsable del tratamiento es la persona física o jurídica, autoridad pública, servicio u otro organismo que trate datos personales por cuenta del responsable del tratamiento. Falsa

El delegado de protección de datos es la persona física o jurídica, autoridad pública, servicio u organismo distinto del interesado, del responsable de tratamiento, del encargado de tratamiento y de las personas autorizadas para tratar los datos personales bajo la autoridad directa del responsable o del encargado.

12. ¿Cuál es la condición básica para poder tratar datos personales del interesado?

a) No hay una condición necesaria.

b) Que el interesado reciba el correcto pago una vez se han usado sus datos.

c) Que el interesado dé su consentimiento.

d) Que el interesado resida en el mismo país en el que se tratan sus datos.

13. ¿Cuáles de los siguientes son derechos de los interesados en relación al tratamiento de sus datos?

- a) Derecho al olvido.
  - b) Derecho a la duplicidad.
  - c) Derecho de acceso.
  - d) Derecho de rectificación.
14. ¿Qué es la seudonimización de datos?
- a) Un mecanismo para evitar que entre los datos acerca de un usuario figure su género.
  - b) Un mecanismo para que todas las fechas de nacimiento de los interesados se eliminen automáticamente.
  - c) Un mecanismo para que cada identificador único en las bases de datos se almacene en codificación hexadecimal.
  - d) Un mecanismo para evitar que se pueda reconocer la identidad de una persona sin utilizar información adicional.
15. ¿Un encargado de tratamiento de datos puede subcontratar a su vez el tratamiento de datos?
- a) No puede en ningún caso.
  - b) Sí, pero únicamente contando con la autorización previa y por escrito del responsable y mediante contrato con el interesado de modo que éste sepa que la ley se está cumpliendo.
  - c) Sí, en cualquier caso.
  - d) Sí, pero únicamente contando con la autorización previa y por escrito del responsable y mediante contrato con el subcontratado de modo que quede obligado a cumplir con sus mismas obligaciones.
16. ¿Cuáles de las siguientes son medidas de seguridad que se deben cumplir al tratar datos personales?
- a) El cifrado de los datos personales.
  - b) La ocultación de fechas de nacimiento.
  - c) La replicación en distintos clusters de toda la información.
  - d) La capacidad de restaurar los datos personales de forma rápida.
17. ¿Cuáles de los siguientes son aspectos que cubre el RGPD?
- a) Transferencias internacionales de datos.
  - b) Supervisión por las autoridades de protección de datos.
  - c) Responsabilidades y sanciones.
  - d) Tipos de impuestos aplicables.
  - e) Códigos de conducta.
18. ¿Cuál es la relación existente entre Gestión de Datos y Gobierno de Datos?
- a) El Gobierno de Datos supervisa la Gestión de Datos.
  - b) El Gobierno de Datos es supervisado por la Gestión de Datos.
  - c) Son equivalentes.
  - d) El Gobierno de Datos es una de las actividades contempladas dentro de la Gestión de Datos.
19. ¿Cuáles de los siguientes están entre los objetivos del Gobierno de Datos?
- a) Asegurar la calidad de los datos.
  - b) Evitar la necesidad de realizar procesamiento por lotes.
  - c) Reducir las necesidades de memoria RAM de los nodos del clúster.

- d) Asegurar la correcta gestión de metadatos.
  - e) Asegurar que el ciclo de vida de los datos es claro y está controlado.
20. ¿Cuáles de las siguientes son actividades en las que está presente el Gobierno de Datos?
- a) Definición del modelo de metadatos, incluyendo tanto su descripción técnica como de negocio.
  - b) Optimización del desempeño de los microprocesadores.
  - c) Políticas y actividades para la documentación de los datos a lo largo de su ciclo de vida.
  - d) Almacenamiento de datos, mecanismos de despliegue y administración de procesos de carga.
  - e) Control de la temperatura alcanzada por los componentes de los nodos.
  - f) Diseño y desarrollo de políticas, estándares, auditoría de la seguridad y cumplimiento regulatorio.
  - g) Políticas y actividades para la documentación de los datos a lo largo de su ciclo de vida.
  - h) Monitorización de las funciones de red.
21. El *data owner* es aquel que ejecuta la gestión de los datos. Falsa
- Los custodios del dato atienden las peticiones de tecnología de los propietarios de los datos.  
 El CDO es responsable de implementar las políticas y procesos definidos.  
 Los *data stewards* son responsables de implementar las políticas y procesos definidos.